PATENT ABSTRACTS OF JAPAN

(11)Publication number : **05-028717**
(43)Date of publication of application : **05.02.1993**

(51)Int.Cl. **G11B 27/024**
G06F 15/62
H04N 5/445
H04N 5/782
H04N 5/93

(21)Application number : **03-181010**
(22)Date of filing : **22.07.1991**

(71)Applicant : **SONY CORP**
(72)Inventor : **OBA AKIO**

(54) **DISPLAY DEVICE**

(57)Abstract:
PURPOSE: To easily recognize the whole flow of a sound by successively detecting and displaying the outline of continuous sound data.
CONSTITUTION: A video signal and an audio signal outputted from a VTR 14 are respectively written in a memory 16V and 16A at each (n) frame. The sampling of the picture data of the memory 16V is operated by an image processor 17and the data are compressedand successively written in a memory 23so that a reduced screen can be prepared An audio processor 30 reads out the sound data from the memory 16A in order to detect a sound leveland extracts the characteristic point of the sound data in order to decide a color. Thenthe data of the color for a length according to the sound level are written at the sound display part at the lower part of each reduced screen of the memory 23corresponding to the picture data of each frame. Thusthe whole flow of the sound can be precisely recognizedand an editing operation can be efficiently operated by using the data with animation data.

CLAIMS

[Claim(s)]
[Claim 1]A display comprising:
A detection means to detect an outline of continuous voice data one by one.
A displaying means which displays the above-mentioned outline detected visually.

DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Industrial Application]This invention is usedfor example for edit of videotapeand relates to a suitable display.
[0002]
[Description of the Prior Art]The check of the outline of a raw material and the election of a desired cut which are recorded as a visible image like what is called SHINEFIRUMU can be performed only by viewing the raw material.
[0003]In order to know the outline of the raw material in the case of the raw material in which dynamic image datasuch as videotape and a video diskis recorded in the invisible state(a) How to reduce the video of two or more frames to the method (b) monitor which displays one screen at a time on a monitorand performs a high-speed search etc. if neededand display on it in scrolling by multi-picture features (refer to JP61-44437B) ** is adopted.
[0004]In the method of (a)checking the outline of the videotape of the TV program of a thing for 1 hourfor example took the time beyond itand there was inconvenience to which editing efficiency worsens. In the method of (b)the short cut of TV commercials etc. might be overlookedfor exampleand there was inconvenience which does not have the reproducibility of the check and has dispersion by a worker.
[0005]Thenthese people displayed as a still picture which made the outline of a series of dynamic image data correspond to time progressand compressed it previouslyand proposed what can recognize the overall flow of a picture with sufficient accuracy (refer to JP2-260075A).
[0006]
[Problem(s) to be Solved by the Invention]By the wayfor example in edit of videotapeif the overall flow of not only the overall flow of a picture but a sound can be recognizedit will become still more convenient.
[0007]Soin this inventionthe display which can grasp an audio overall flow easily is provided.
[0008]
[Means for Solving the Problem]This invention is characterized by comprising:
A detection means to detect an outline of continuous voice data one by one.
A displaying means which displays an outline detected visually.

[0009]
[Function]The outline of voice datai.e.a sound levela kindetc. make a displaying means correspond to time progressand are continuously displayed on it. Thereforean audio overall flow can be recognized now with sufficient accuracy.
[0010]
[Example]Hereafterone example of this invention is describedreferring to drawings.
[0011]Drawing 2 shows processing of the dynamic image data twisted to this example. In the figure1 shows the video picture group as a whole. It is possible that this video picture group 1 put in order the video corresponding to a series of dynamic image data currently recordedfor example on videotapea video disketc. in the direction of time (t axis) in the unit of the frame 2.
[0012]Since the frame frequency of a video signal is 30 Hz (NTSC system)the 30 frames

2 are allotted in 1 second as the video picture group 1. 2A-2E show a series of frames of the video picture group 1.

[0013]4 is a vertical slit for an image data input (input slit) set up on the frame 2and the picture of the frame 2 is sampled by this vertical slit 4. This vertical slit 4 is scanned with a prescribed speed horizontally (the direction of H)and if it arrives at the right end section of the frame 2it is again scanned repeatedly in the direction of H from a left edge part. Thereforethe vertical slit 4 is scanned in the slanting direction of Ht in three-dimensional space including a time-axis.

[0014]When the vertical slit 4 scans the video picture group 1 from a left edge part to a right end section in the direction of Htthe f frames 2 are crossed and it is assumed that it is that by which one slit shape picture is sampled by this vertical slit 4 about n frames (usually n= 1).

[0015]If f is chosen as the multiple of nusing the predetermined integer Xf=nX will be materialized and the slit shape picture of X individual will be sampledrespectively from the frame groups (3A3Betc.) which consist of the f frames 2.

[0016]In this examplethe time taken for the vertical slit 4 to scan the video picture group 1 from a left edge part to a right end section in the direction of Ht is set as 12 secondsand it is set as n= 1 and set to f=X=12x30=360.

[0017]And after the slit shape picture of X individual obtained from the frame group 3A is connected horizontallyand it is compressed perpendicularlyrespectively and the reduced screen 6A is formed in it. The reduced screen 6A is inserted in into the display screen 5 corresponding to one frame memory. It is compressed after similarly the slit shape picture of X individual obtained from the frame group 3B is connected horizontallyand the reduced screen 6B is formedand this reduced screen 6B is inserted in next to the reduced screen 6A in the display screen 5.

[0018]Even if it corresponds to the following frame groupsa reduced screen is formed similarlyand it is inserted in one by one into the display screen 5.

[0019]Actuallysince one slit shape picture sampled by the vertical slit 4 is serially generated at a timeit is compressed in order of generation and inserted in at a time into [one] the display screen 5.

[0020]For exampleit is assigned so that the vertical slits 4A-4E may scan sequentially in the direction of H corresponding to the frames 2A-2Erespectively. The picture on the slit sampled by each vertical slits 4A-4E is compressedrespectivelyand let it be a picture of the portion of the vertical slits (output slit) 7A-7E of the display screen 5.

[0021]Methods of compressing the picture sampled by the vertical slits 4A-4E include the method of only thinning out image dataand the method of taking the weighted average of a predetermined region. When only thinning out image datathough the slits 4A-4E have the width for 1 pixel in the direction of Hthey are good.

[0022]The phonological representation part 60 is formed in the lower part of each reduced screen of the display screen 5. Corresponding to each of the slit shape picture which constitutes each reduced screenthe sound level Ea corresponding to the picture is displayed on each phonological representation part 60 so that an enlarged display may be carried out to drawing 3. And the display of the phonological representation part 60 corresponding to each reduced screen is performed with the kind of sound corresponding to each reduced screenfor examplea voicemusicand the other colors to which it responded.

[0023]Drawing 1 shows the display of this example. In the figure8 shows the host

computer. This host computer 8 functions as a control means of the whole device.
[0024]11 is a system bath12 is a keyboardand an operator enables it to give various commands to the host computer 8 via the input output circuit 13 and the system bath 11 from the keyboard 12.
[0025]A Video RAM (VRAM1) and 16A of VTR as a source of dynamic image dataand 15V and 15A are the memories (ARAM) for audio signals an A/D converter and 16V 14. The video signal for one frame is memorized by the memory 16Vand the audio signal for one frame is memorized by the memory 16A.
[0026]The video signal (for examplecomponent signal of YR-YB-Y or RGand B) outputted from VTR14 is written in back Video RAM 16V changed into the digital data with A/D converter 15V. The audio signal outputted from VTR14 is written in the back memory 16A changed into the digital data with A/D converter 15A.
[0027]Corresponding to a actual display screenas shown in drawing 4it is considered as HL dot horizontally (VX1 direction)and let the storage area of Video RAM 16V (VRAM1) be VL dot perpendicularly (VY1 direction). The address of each picture element data read from this Video RAM 16V is directed with coordinates (VX1VY1) (0 <=VX1 <=HL-10 <=VY1 <=VL-1).
[0028]17 is an image processor. At the time of video-index creationby the image processor 17. The data of the portion surrounded by the vertical slit 4 (refer to drawing 4) from the image data for one frame of Video RAM 16V is readIt is written in the portion surrounded by the vertical slit (output slit) 7 (refer to drawing 5) of Video RAM (VRAM2) 23 which the data is compressed and consists of frame memories. In additionthe image processor 17 has a function which displays directions cursor on the screen corresponding to Video RAM 23.
[0029]If the image processor 17 is expressed as a set of the means corresponding to a functionthis image processor 17 will consist of the input slit transportation device 18the slit data reading means 19the output slit transportation device 20the slit data writing means 21and the directions cursor display means 22.
[0030]Corresponding to a actual screenas shown in drawing 5it is considered as HL dot horizontally (VX 2-way)and let the storage area of Video RAM 23 (VRAM2) as well as Video RAM 16V be VL dot perpendicularly (VY 2-way). The address of each picture element data written in this Video RAM 23 is directed with coordinates (VX2VY2).
[0031]30 is an audio processor. At the time of audio index creationintegration treatment of the voice data for one frame written in one by one is read and carried out to the memory 16A by the audio processor 30and the sound level Ea is detected. And color data is written in the slit region of the phonological representation part 60 of the lower part of the vertical slit 7 of Video RAM 23 by the length corresponding to the sound level Ea.
[0032]In the period corresponding to one reduced screenaudio information is read from the memory 16Afor examplethe focus is extracted by the neural network. The kind of sound of a voicemusicand others is distinguished by thisand a color is chosen from a color map according to that kindand let the data of this color be color data written in the phonological representation part 60 of the lower part of each reduced screen as mentioned above.
[0033]If the audio processor 30 is expressed as a set of the means corresponding to a functionthis audio processor 30 will consist of the audio information reading means 31the characteristic point extracting means 32the color determining means 33the level detection

means 34and the color data writing means 35.
[0034]24 is RAM for cursor for memorizing the data of cursor. The data of the cursor read from RAM24 for picture element data and cursor read from Video RAM 23 is supplied to the synthetic circuit 25and composite image data is formed.
[0035]The composite image data outputted from the synthetic circuit 25 is changed into an analog signal with D/A converter 26and is supplied to the monitor 27 or a video printer (not shown)and it is supplied also to the external storages (VTRa floppy disketc.) 28.
[0036]It enables it to write the video signal reproduced from the external storage 28 in Video RAM 23 via A/D converter 29 and the system bath 11.
[0037]At the time of index creationthe image data of the video picture group 1 outputted from VTR14It is written in Video RAM 23 (it illustrates to drawing 5) as a series of slit data via Video RAM 16V (it illustrates to drawing 4)A series of operations at the time of the color data in which the sound level Ea and a kind are furthermore shown at Video RAM 23 based on audio information being written in are explained for every step along with the flow chart of drawing 6.
[0038]In this casewhat compressed collectively X slit data which it extracted one piece at a time from each frame of Video RAM 16V shall be written in as the reduction images 6A and 6B which consist of a pixel of the individual (XxY) of Video RAM 23and ...
[0039][Step 101] **X and **Y are calculated according to the following formula.
[0040]As for **X=HL/X**Y=VL/Y**X and **Yit is good not to be an integerand they compress into the value of the one pixel 34 of Video RAM 23 the picture element data of the block 33 which consists of a pixel of the individual (**Xx**Y) of Video RAM 16V.
[0041]In this examplein order to compress simplylet the data of the pixel which makes an address the coordinates (VX1VY1) of the upper left corner of the block 53 of Video RAM 16V be data of the pixel 54 which makes an address the coordinates (VX2VY2) of Video RAM 23 as it is. Since coordinates (VX1VY1) stop being integral pairs when nonintegral**X and **Y calculate the value which is a pixel which coordinates (VX1VY1) direct with the interpolation from the value of the surrounding pixel.
[0042]The horizontal arrangement number h of the reduced screens (6A6Betc.) which consist of the vertical slit 7 of X individual in Video RAM 23 is calculated according to the following formula. Xs is a pixel number of horizontal space.
[0043]h=(HL-Xs)/X -- again -- The reduced screens 6A and 6B ... number FR -- respectively -- 0 and 1 -- it is referred to as ... and FR0 and initializes to FR=0.
[0044]The image data for one frame outputted from VTR14 is written in Video RAM 16Vand the audio information outputted by corresponding from VTR14 is written in the memory 16A.
[0045][Step 102] From the memory 16Aaudio information is read and extraction and color decision processing of the focus are started by the audio processor 30.
[0046][Step 103] It calculates by following BX in the coordinates of the upper left corner of the reduced screens (6A6Betc.) of number FR of Video RAM 23and following the following formula in BY as (BXBY). About the pixel number of left end spaceYs1 is a pixel number of the space of a vertical upper bed Xs1. It is YA =Y+Y'+Y"and Y' is a pixel number of the perpendicular direction of the phonological representation part 60and Y " is a pixel number of the perpendicular direction between the phonological representation part 60 and a reduced screen (refer to drawing 5).

[0047]BX=(FRmod h) X+Xs1BY=[FR/h] YA+Ys1 -- in the formula of these(FRmod h) shows remainder of FR/h and [FR/h] shows the maximum integer that does not exceed FR/h.
[0048][Step 104] The initial value of coordinates VX1 of the vertical slit 4 of Video RAM 16V and coordinates VX2 of the vertical slit 7 of Video RAM 23 is set as 0 and BXrespectively.
[0049][Step 105] The initial value of coordinates VY1 of the vertical slit 4 of Video RAM 16V and coordinates VY2 of the vertical slit 7 of Video RAM 23 is set as 0 and BYrespectively. It initializes to N= 0.
[0050][Step 106] Only 1 increases the value of N.
[0051][Step 107108] the image processor 17After reading the data of the pixel of the coordinates (VX1VY1) of Video RAM 16V and writing in as data of the pixel of the coordinates (VX2VY2) of Video RAM 23only **Y increases the value of coordinates VY1 and only 1 increases the value of coordinates VY2.
[0052][Step 109] When the data of the vertical slit 4 of Video RAM 16V is read in the D1 directionthe data of the vertical slit 7 of Video RAM 23 is written in D 2-way. And when coordinates VY1 of the vertical slit 4 of Video RAM 16V is below VLit returns to Step 106and when coordinates VY1 exceeds VLit progresses to Step 110.
[0053][Step 110] The audio processor 30 reads and integrates with all or a part of audio information of 1 frame period currently written in the memory 16A.
[0054][Step 111] It is A= [DI/DMxY']when setting the integrated output corresponding to the pixel number of the perpendicular direction of the phonological representation part 60 of Video RAM 23 to DM and setting an integrated output to DI.
It calculates. [DI/DMxY'] is the maximum integer that does not exceed DI/DMxY'. The data of this A shows the sound level Ea.
[0055][Step 112]
It is set as DXN =VX2DYN =VY2+Y'-ADYN '=VY2+Y'-1.
[0056][Step 113114] Only **X increases the value of VX1 and only 1 increases the value of VX2. This means that only **X moves the position of the vertical slit 4 of Video RAM 16V to the rightand only 1 moves the position of the vertical slit 7 of Video RAM 23 to the right.
[0057]And from VTR14the host computer 8 inputs the image data and audio information of a frame of the n-th sheet from the present frameand writes them in Video RAM 16V and the memory 16Arespectively.
[0058][Step 115] When coordinates VX1 of the vertical slit 4 of Video RAM 16V is below HLit returns to Step 105. Since it means that one scan to the horizontal direction of the vertical slit 4 of Video RAM 16V was completed when coordinates VX1 exceeds HLit progresses to Step 116.
[0059][Step 116] The color data of the color determined by the coordinates (DXNDYN) -(DXNDYN ') the audio processor 30 of Video RAM 23 is written in. Hereit is N=1-X. By thiscorresponding to the image data of each slita sound level will be shown in the phonological representation part 60 of a Video RAMand the color data in which the kind is shown will be written in.
[0060][Step 117118] Number FR0 of the reduced screen which 1's increases number FR of a reduced screen and Video RAM 23 permits [the number FR] When it is the followingit returns to Step 102. When number FR exceeds number FR0it means that

creation of the video index for one screen was completed. Thereforeit progresses to Step 119 and post-processing is carried out.
[0061]It is possible to accumulate the image data of Video RAM 23 in the external storage 28 via D/A converter 26 as post-processingor to supply the image data to the monitor 27 via D/A converter 26.
[0062]The video picture group 1 which is a set of dynamic image data will be accumulated in the external storage 28 in the form which carried out the data compression by the method of expressing it also as a kind of video slice (it illustrates to drawing 7). A sound level will be shown corresponding to two or more reduced screen and each reduced screenand the index which shows the kind will be displayed on the monitor 27.
[0063]Thenit will return to Step 101 again by operation of an operatorand the index to the video signal and audio signal of a continuation from VTR14 will be created.
[0064]The index created as mentioned above is the reduced screen which compressed and connected the image data which changes the position of the vertical slit 4 and is sampled for every n frame of the video picture group 1.
The outline of the dynamic image data of the video picture group 1 is made to correspond to time progressand can be checked.

[0065]In order to make it change here so that the position of the vertical slit 4 may be scanned from the left end of one screen to a right end in 12 secondswhen forming the N reduction images 6A and 6B and ... into the display screen 5it is 12xN=12N [a second]. 30x12xN=360N [a frame]
Morethe dynamic image data of the video signal for 12 N seconds (360N frame) will be compressed and displayed into the display screen 5 of one sheet (frame)and a series of dynamic image data is compressed with the very big compression ratio.
[0066]And since the vertical slit 4 is scanning from the left end to the right end when the video of the video picture group 1 changes slowly on the basis of 12 secondsthe state of the original picture of abbreviated video can be restored. On the other handif the video of the video picture group 1 changes rapidly like commercialsthe open circuit which changes to any of a reduced screen they are nonsequetially will be formed. Thereforein spite of compressing and displaying dynamic image data with the compression ratio of abbreviated 1/360Nthe outline of a portion in which it is changing slowly can be checkedand the portion which changes rapidly can be checked as an open circuit.
[0067]Corresponding to the slit image which constitutes each reduced screena sound level is displayed on the phonological representation part 60. That isa sound level is displayed corresponding to time progress. The color according to an audio kind is given to the display of this sound level for every reduced screen. Thereforean audio overall flow can be easily grasped with a pictureand editing work etc. can be performed much more efficiently.
[0068]Instead of sampling image data using the vertical slit 4 from each frame 2 of the video picture group 1horizontal slits may be used and image data may be sampled. In this caseit is constituted so that horizontal slits may be periodically scanned from the upper bed of the frame 2 to a lower end perpendicularly (the direction of V) with a prescribed speed. And the phonological representation part 60 will be formed in the left of a reduced screenor a right flank.

[0069]Although an audio kind is displayed by changing a colorit may be made to display by changing the size and luminosity of a viewing area in the above-mentioned example.
[0070]
[Effect of the Invention]In this inventionthe outline of voice datasuch as an audio level and a kindmakes a displaying means correspond to time progressand is continuously displayed on it.
Thereforean audio overall flow can be recognized with sufficient accuracy.
Thereforeediting work etc. can be performed much more efficiently by using itfor example with the outline of dynamic image data.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]
[Drawing 1]It is a lineblock diagram of an example.
[Drawing 2]It is a figure showing the concept of an index.
[Drawing 3]It is the figure which carried out the enlarged display of the phonological representation part.
[Drawing 4]It is a figure showing the data structure of a Video RAM.
[Drawing 5]the data structure of a Video RAM is shown -- it comes out.
[Drawing 6]It is a flow chart which shows the operation at the time of index creation.
[Drawing 7]It is a figure showing the relation between an input video signal and a video index.
[Description of Notations]
1 Video picture group
2 Frame
3A3B frame group
4 Vertical slit (input slit)
5 Display screen
6A and 6B Reduced screen
7 Vertical slit (output slit)
8 Host computer
14 VTR
16A The memory for audio information
16V23 Video RAMs
17 Image processor
27 Monitor
30 Audio processor
60 Phonological representation part